Title: Transformer-Based Atmospheric River Segmentation Model from Climate Reanalysis Data

Authors: Seo Yen Lim[1]*, So Yeon Kim[1]*, Ashesh Chattopadhyay[2], Sung Beom Cho[3], Hyunsouk Cho[3], and Soo Kyung Kim[1+]

1. Ewha Womans University 2. University of California Santa Cruz 3. Ajou University
(* Equal Contribution, + Corresponding Author)

Abstract

Atmospheric rivers (ARs)—long, narrow corridors of water vapor transport—play a central role in the global hydrological cycle and drive extreme precipitation events that cause floods and socio-economic losses. Traditional AR detection methods rely on threshold-based heuristics applied to reanalysis variables such as integrated vapor transport (IVT) and sea-level pressure. However, these methods are sensitive to threshold choices, struggle with irregular AR geometry, and produce inconsistent climatologies across datasets. Recent advances in deep learning (DL) have demonstrated the potential for more robust AR detection.

In this study, we evaluate transformer-based segmentation architectures for AR detection using multi-channel ClimateNet reanalysis data. Specifically, we adapt **MaskFormer**, which reformulates segmentation as mask-level classification, and **ChannelViT–U-Net**, a channel-wise vision transformer combined with a lightweight decoder. Both models are trained on ERA5-derived ClimateNet samples using four meteorological variables (U850, V850, PSL, TMQ) and evaluated against expert-labeled AR masks. For comparison, we also benchmark against a convolutional baseline (U-Net with MobileNetV3 encoder).

Our experiments demonstrate that transformer-based models outperform the CNN baseline in capturing the elongated, irregular geometry of ARs, with MaskFormer achieving the highest AR Intersection-over-Union (IoU) across test scenarios. We further analyze the effects of patch size, class imbalance weighting, and decoder design choices. Results suggest that geometry-aware segmentation models such as MaskFormer provide more reliable AR boundaries, even with limited training data. Qualitative visualizations show that transformer-based models better capture coherent AR structures, highlighting the limitations of pixel-wise CNN predictions.

This work underscores the promise of transformer architectures for extreme weather detection, while emphasizing challenges such as rare-event imbalance, patch-resolution trade-offs, and evaluation metrics beyond mean IoU. Our findings point toward future directions integrating multi-scale features, physical constraints, and richer datasets to develop interpretable, robust AR monitoring systems.

References

- Bao, Yujia, Srinivasan Sivanandan, and Theofanis Karaletsos. *Channel Vision Transformers: An Image Is Worth 1 × 16 × 16 Words*. arXiv preprint arXiv:2309.16108, 2024.
- Cheng, Bowen, Alexander G. Schwing, and Alexander Kirillov. *Per-Pixel Classification Is Not All You Need for Semantic Segmentation*. arXiv preprint arXiv:2107.06278, 2021.
- Guan, Bin, and Duane E. Waliser. "Detection of Atmospheric Rivers: Evaluation and Application of an Algorithm for Global Studies." *Journal of Geophysical Research: Atmospheres* 120, no. 24 (2015): 12,514–12,535.
- Prabhat, Kashinath, Mudigonda, S., Kim, S., Kapp-Schwoerer, L., Graubner, A., et al. "ClimateNet: An Expert-Labeled Open Dataset and Deep Learning Architecture for Enabling High-Precision Analyses of Extreme Weather." *Geoscientific Model Development* 14, no. 1 (2021): 107–124.
- Shields, Christine A., et al. "Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Project Goals and Experimental Design." *Geoscientific Model Development* 11, no. 6 (2018): 2455–2474.